Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/017237

International filing date: 01 June 2004 (01.06.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/475,069
Filing date: 30 May 2003 (30.05.2003)

Date of receipt at the International Bureau: 19 August 2004 (19.08.2004)
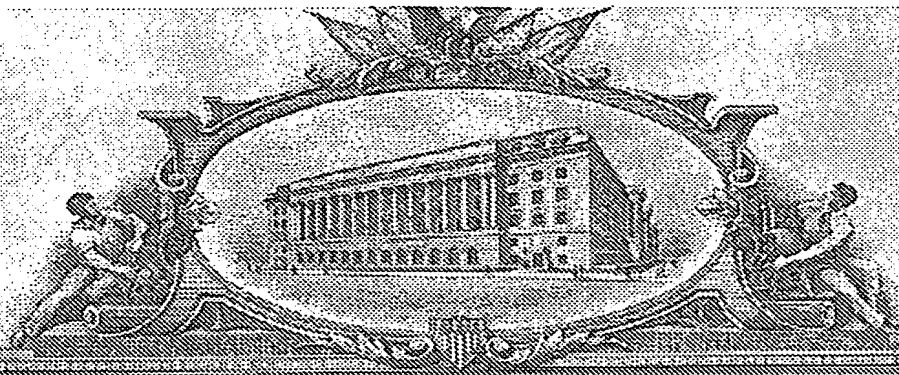
Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1209035

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 10, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/475,069*
FILING DATE: *May 30, 2003*
RELATED PCT APPLICATION NUMBER: *PCT/US04/17237*

Certified by

Jon W Dudas

Acting Under Secretary of Commerce
for Intellectual Property
and Acting Director of the U.S.
Patent and Trademark Office

06/03/03   60475069 . 05/30/03

PROVISIONAL APPLICATION FOR PATENT COVER SHEET
This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR §1.53(c).

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Chung-kuan	Cheng	San Diego, California
Zhengyong	Zhu	La Jolla, California

Additional inventors are being named on the __0__ separately numbered sheets attached hereto

TITLE OF THE INVENTION (500 characters max)
Circuit Network Analysis Using Adaptive Algebraic Multigrid Approach

CORRESPONDENCE ADDRESS

Direct all correspondence to:

[X] Customer Number: 20985

OR

[ ] Firm or Individual Name	
Address	
Address	
City		State		ZIP	
Country	United States	Telephone		Fax	

ENCLOSED APPLICATION PARTS (check all that apply)

[X] Specification   *Number of Pages*   13   [] CD(s), Number [   ]

[] Drawing(s)   *Number of Sheets* [   ]   [] Other (specify) [   ]

[] Application Data Sheet. See 37 CFR 1.76.

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

[X] Applicant Claims small entity status. See 37 CFR 1.27.   FILING FEE

[X] A check or money order is enclosed to cover the filing fees.   AMOUNT ($)

[] The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: [06-1050] $80

[] Payment by credit card. Form PTO-2038 is attached.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
[X] No.
[] Yes, the name of the U.S. Government agency and the Government contract number are:

Respectfully submitted,

Signature _____   Date _____ May 30, 2003 _____

Typed Name Bing Ai, Reg. No. 43,312 _____

Telephone No. (858) 678-5070 _____

Docket No. 15670-029P01 _____

10289348.doc

PROVISIONAL APPLICATION FOR PATENT

under

37 CFR §1.53(c)

TITLE: CIRCUIT NETWORK ANALYSIS USING ADAPTIVE
ALGEBRAIC MULTIGRID APPROACH

APPLICANT: CHUNG-KUAN CHENG AND ZHENGYONG ZHU

Circuit Network Analysis Using Adaptive Algebraic Multigrid Approach

[0001]   This application relates to analysis of circuit

5   networks such as power networks and clock networks. This

application includes two Attachments I and II on pages 4-13 as

part of this application.

[0002]   The network analysis methods described here may be

generally categorized as algebraic multigrid (AMG) methods. An

10   AGM method can be used to handle circuit networks with irregular

structures. The present network analysis methods significantly

improve the capabilities of the conventional AGM methods in

various aspects. For example, adaptive grid structures and

adaptive smoothing operations are used to capture the multi-rate

15   behavior and circuit latency in various networks. The circuit

spatial and temporal latency are fully explored along with the

mutli-rate behavior to avoid any unnecessary computation without

losing accuracy and convergence of the analysis. The present

methods also have the ability to analyze coupling effects with

20   electromagnetic retardations for high-frequency chi-package-

board scale analysis. Hence, as one example, the effects of

inductance in the circuits can be included in the analysis.

[0003]   More detailed description is included in Attachments I

and II, where exemplary implementations in Attachment I on pages

25   4-7 are referred to as "Adaptive AMG," and exemplary

implementations in Attachment II on pages 8-13 are referred to

as "SPICE$_{Diego}$." Several exemplary implementations of the network

analysis methods described here are summarized as follows.

5   1. A method for analyzing a circuit network, comprising:

representing a circuit network by using a matrix of nodes

having fine nodes and coarse nodes;

applying an adaptive coarse grid construction procedure to

assign grid nodes in the matrix as either coarse grid nodes or

10   fine grid nodes according to (1) circuit activities and (2) a

matrix structure of the matrix, where a node with a maximum

potential is assigned as a coarse node and an unassigned node

adjacent to a coarse node is assigned as a fine node; and

applying iterative smoothing operations at selected local

15   fine grids corresponding to active regions at a finest level

obtained in the adaptive coarse grid construction procedure.

2. The method as in 1, wherein coarse grid nodes are

divided into non-adaptive coarse nodes which are selected

20   according to the matrix and adaptive coarse nodes which are

selected according to circuit activities.

3. The method as in 2, wherein, in assigning non-adaptive

coarse nodes, a potential of each neighboring node of a newly

assigned fine node is increased by one unit before a next level

of assigning coarse and fine grid nodes so that each fine node

has at least one neighboring coarse node.

5   4. The method as in 2, wherein an adaptive coarse node I

.   selected according to a first-order derivative of a nodal

voltage.

10263136.doc

10

ATTACHMENT I

Power Network Analysis Using an Adaptive Algebraic Multigrid Approach

ABSTRACT
In this paper, we introduce an efficient analysis method for the power network of general topology. The new approach is based on algebraic multigrid (AMG) method that can avoid the slow convergence of basic iterative methods. An innovative adaptive coarsening and error-smoothing scheme is employed to further speed up the performance, taking advantage of the spatial variation of power supply noise. Experimental results show that our method is more than 100 times faster than SPICE3.

Categories and Subject Descriptors
B.7.2 [Design Aids]: Simulation

General Terms
Algorithms, Performance, Experimentation

Keywords
Algebraic multigrid method, Adaptive analysis, Power distribution network, Circuit Simulation

1. INTRODUCTION
Power network analysis covers voltage drop, voltage oscillation, and electromigration [1,2]. Excessive voltage drops reduce the switching speed as well as the noise margins of circuits and cause logic failures. Electromigration can decrease the chip lifetime [3]. Previous on-chip power grid analysis focuses on IR-drop caused by the resistance of power network. With the signal frequency increasing rapidly, the $j\omega L$ of on-chip power network becomes comparable with the resistance. Hence the inductance effect cannot be ignored. Moreover voltage oscillation may occur when power network resonance frequency drops to the range of the signal frequency [5].

One bottleneck of power network analysis is the tremendous amount of elements in power network. Direct methods such as LU decomposition used in SPICE are thus prohibitive. To address this bottleneck, M. Zhao and Y. Cao used hierarchical macro modeling or model order reduction techniques in [6][7]; H. Chen divided the chip into 100*100 equipotential segments to reduce the complexity in [8]; Y. Lee adopted the ADI concepts in [9].

Some sparse matrix techniques have also been applied to the power network simulation. T. Chen employed a Preconditioned Conjugate Gradient (PCG) solver in [10,18] and Nassif developed a multigrid method in [11][12].

For typical power grid, the RLC values are quite uniform at the same layer. But due to the non-uniform power densities and the timing of the switching events, power supply noise exhibits spatial variation, which means some nodes of the power network have more rapid nodal voltage changes, or in other words, these nodes are more active than the rest circuit.

In this paper, we present an adaptive approach for power grid analysis based on algebraic multigrid (AMG) method [15]. There are two kinds of multigrid methods: geometric multigrid and algebraic multigrid. Geometric multigrid requires regular mesh structure but has less coarse grid reduction cost while algebraic multigrid (AMG) can handle problems in general topology but requires heavier coarsening overhead [15].

Adaptive concepts can be integrated with the coarsening scheme and smoothing operation in multigrid method. Some adaptive methods like FAC [19] and MLAT [20] can be applied to geometric multigrid method. The basic idea behind these adaptive methods is that active regions should have finer grid structure, since active subnetworks need more computation to model their behavior accurately.

The main differences between our approach and the multigrid method in [11][12] are

1) Our approach covers self and mutual inductances, while [11][12] can only handle RC networks. This is due to the limitation of geometric multigrid coarsening algorithm used in [11][12]. Since multigrid method requires the system matrix to be symmetric positive definite (S.P.D.) which circuits with inductors cannot satisfy using Modified Nodal Analysis method (MNA), reformulation as discussed in section 2 is needed. But the topology of reformulated system matrix is no longer the same as original circuit topology. It is not easy to derive coarse level grid directly from circuit topology in this kind of situation even if the power network is a regular mesh.

2) Our coarsening scheme makes no assumption about the topology of power network. While in [11][12], special consideration is needed for irregular power networks.

3) [11][12] ignored the error smoothing operation and only executed one multigrid iteration cycle. Although it may not result in much error for well-designed power network, accuracy cannot always be guaranteed for general cases. We keep the smoothing operations and iterate the multigrid cycles until the norm of the residue is less than a threshold.

4) We adopt an adaptive coarsening and smoothing scheme, which makes it possible to assign more computation to the active subnetworks and reduce the unnecessary computation for inactive subnetworks.

2. OVERVIEW

We focus on power source network with the assumption that power and ground can be separated. Usually power network is an irregular mesh as shown in Figure 2.1. Every intersection node has a ground capacitance. Between neighboring intersection nodes is resistor or resistor and inductor in series. Mutual inductance can also be included. The devices with activities are modeled as time-varying current sources [8].

When only RC model is included, system equation can be formulated as

$$C\dot{X}(t) + GX(t) = U(t), \qquad (2.1)$$

where X is the vector of nodal voltages.
Appling trapezoidal approximation with time step h to equation (2.1), we have

$$(G + \frac{2}{h}C)X(t+h) = -(G - \frac{2}{h}C)X(t) + U(t) + U(t+h) \qquad (2.2)$$

The left hand side matrix in (2.2) is symmetric and positive definite which makes the iterative methods converge quickly.
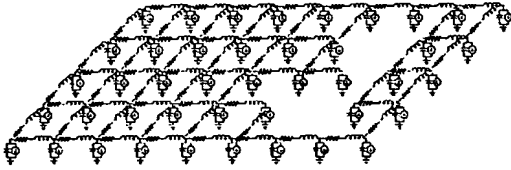

Figure 2.1 Power Network Structure

When inductance is included, the system matrix equation becomes:

$$\hat{C}\dot{X}(t) + \hat{G}\,\hat{X}(t) = \hat{U}(t) \qquad (2.3)$$

where $\hat{C} = \begin{bmatrix} C & 0 \\ 0 & L \end{bmatrix}$, $\hat{X} = \begin{bmatrix} V \\ I \end{bmatrix}$, $\hat{U} = \begin{bmatrix} U \\ 0 \end{bmatrix}$ and $\hat{G} = \begin{bmatrix} G & -A_l^T \\ A_l & 0 \end{bmatrix}$

Equation (2.3) can be rewritten as:

$$\begin{bmatrix} C & 0 \\ 0 & L \end{bmatrix}\begin{bmatrix} \dot{V} \\ \dot{I} \end{bmatrix} + \begin{bmatrix} G & -A_l^T \\ A_l & 0 \end{bmatrix}\begin{bmatrix} V \\ I \end{bmatrix} = \begin{bmatrix} U \\ 0 \end{bmatrix} \qquad (2.4)$$

Appling trapezoidal approximation with time step h, the solution of equation (2.4) is derived by

$$\begin{bmatrix} \frac{2C}{h}+G & -A_l^T \\ A_l & \frac{2L}{h} \end{bmatrix}\begin{bmatrix} V(t+h) \\ I(t+h) \end{bmatrix} = \begin{bmatrix} \frac{2C}{h}-G & A_l^T \\ -A_l & \frac{2L}{h} \end{bmatrix}\begin{bmatrix} V(t) \\ I(t) \end{bmatrix} + \begin{bmatrix} U(t+h)+U(t) \\ 0 \end{bmatrix} \qquad (2.5)$$

Although MNA method can handle elements without admittance description, the transient analysis system matrix in (2.5) is no longer symmetric and positive definite when the inductance is included because of the introduction of current variables.

Since Multigrid as well as PCG method requires the matrix to be symmetric positive definite, some extra processing is needed to reformulate the system [16]. Similar to the method used in [10], we split the variable vector into nodal voltage vector and branch current vector. Using block matrix operations, we decompose (2.5) into two iteration formulas (2.6) for nodal voltages and branch currents, respectively.

$$\Rightarrow \begin{cases} (\frac{2C}{h}+G-\frac{h}{2}A_l^T L^{-1}A_l)V(t+h) = & (2.6) \\ (\frac{2C}{h}-G+\frac{h}{2}A_l^T L^{-1}A_l)V(t)+[U(t)+U(t+h)]+2A_l^T I(T) \\ \quad I(t+h)=I(t)-\frac{h}{2}L^{-1}A_l[V(t+h)+V(t)] \end{cases}$$

In equation (2.6), L^{-1} corresponds to the K matrix [17], where the matrix inversion overhead is reduced by sparsification methods.

If the inductance matrix is symmetric and positive definite (S.P.D.), we can prove that the system matrix is still S.P.D. This also holds for forward Euler, backward Euler integration approximation methods. Note that the topology of reformulated system matrix in (2.6) is no longer the same as original circuit topology. Geometric based coarse grid reduction algorithm cannot be applied to RLC network directly.

3. MULTIGRID METHOD

The multigrid method is an efficient technique first widely used for solving partial differential equations [15]. The basic idea of multigrid method is to map the hard-to-damp low frequency error at fine level to easy-to-damp high frequency error at coarse level, solve the mapped problem at coarse level, and then map the error correction of coarse level back to fine level. The mapping operator from fine to coarse level is called restriction operator I_h^{2h}, and the mapping operator from coarse to fine level is called interpolation operator I_{2h}^h. Here the subscription represents the level. At each level, high frequency error is erased by a smoothing operator that is a forward iterative method such as Gauss-Seidel.
The construction of hierarchical grid structure stops at the level when the reduced matrix can be quickly solved by LU decomposition. Once we find the exact solution of the problem at the coarsest level, we perform the interpolation from the coarsest level back to the find grid. An iteration of restriction from the fine grid to the coarsest grid and interpolation from the coarsest grid back to the fine grid is called a single multigrid cycle, which will iterate several times till converge.

3.1 Geometric Multigrid vs. Algebraic Multigrid

There are two kinds of multigrid methods: geometric multigrid and algebraic multigrid (AMG). Geometric multigrid method requires regular mesh structure. Its coarse grid reduction algorithm is defined geometrically, e.g. coarse node is selected every other fine node.

If we want to handle problems with irregular structure, AMG is a good alternative of geometric multigrid method. The coarsening and interpolation operation of AMG are based on matrix itself. This overhead makes AMG less efficient than geometric multigrid if the problem analyzed has regular mesh structure.

In geometric multigrid, the smoothing operation at each level is important to the convergence. The error after smoothing should be geometrically smooth, in other words, the local error among adjacent nodes should be small. While in AMG, the interpolation plays the crucial role. The convergence rate actually depends on the coarse grid construction and inter-level mapping operators [14].

3.2 AMG Interpolation & Restriction Operator

Since AMG has no grid concept in mind, the inter-level mapping operators have to be determined based on the matrix. In AMG, smooth error means error components with relatively small residuals (3.1) [15], which means after several iterations of smoothing operation, the residue is small but the error decreases very slowly.

$$Ae \approx 0 \qquad (3.1)$$

Equation (3.1) can be rewritten as (3.2). Hence the error of fine node can be well represented by linear combinations of its neighbors' error.

$$a_{ii}e_i \approx -\sum_{j \neq i} a_{ij}e_j \qquad (3.2)$$

If coarse and fine nodes have already been defined, we can further approximate the error of fine node by errors of only its coarse node neighbors. Although there are many other complicated interpolation methods, we find that this simple approach works well for the power grid problem. From (3.2), we can construct the interpolation operator I_{2h}^h. Because of the symmetry of the mapping between levels in two directions, the restriction operator I_h^{2h} is the transposition of the interpolation operator, i.e. $I_{2h}^{h^T} = I_h^{2h}$. The coarse level matrix can be computed as $A^{2h} = I_{2h}^h A^h I_h^{2h}$. Consequently, coarse level matrix A^{2h} remains symmetric positive definite.

If the matrix is symmetric and positive definite, convergence of AMG can be guaranteed as long as the smoothing operation converges at each level. A detailed proof can be found in [13].

4. Adaptive ALGEBRAIC MULTIGRID

In this section, we discuss the adaptive coarse grid construction procedure and the adaptive smoothing operation.

4.1 Adaptive Coarse Grid

In our adaptive coarsening scheme, active regions have relatively finer grid at coarse level. The coarse grid nodes are determined adaptively by the circuit activities as well as by the matrix structure. The coarse grid consists of two kinds of nodes: non-adaptive and adaptive coarse nodes. Non-adaptive coarse nodes are selected by coloring scheme [15] according to the matrix. Adaptive coarse nodes are determined by the circuit activities.

4.1.1 Coloring Scheme for Non-Adaptive Coarse Nodes

A two-level coloring scheme is explained below for sake of clarity. In practice, the coloring scheme is executed at every level except the coarsest one.

Initially, the potential of each node is set to its degree. The node with the maximum potential is selected to be a coarse node and all its unassigned neighboring nodes are set as fine nodes. For each newly set fine node, we increase the potential of its neighbor by 1. The process is repeated until every node has been assigned as fine or coarse node. At the end, each fine node has at least one neighboring coarse node.

4.1.2 Adaptive Coarse Nodes Selection

Adaptive coarse nodes are selected according to the activities of circuits. One good candidate for the measurement of the impact of circuit activities is the first order derivative of nodal voltages that can be approximated from equation (2.4) for RLC network.

For RLC networks as shown in Fig 2.1, since not every node has a ground capacitor, we split voltage vector V into two separate voltage vectors V_1 and V_2, where V_1 is the set of nodes with ground capacitor and V_2 is the set of nodes between resistance and inductance in branches. We then calculate the first order derivative of nodal voltage in V_1.

Rewriting equation (2.4), we have

$$\begin{bmatrix} C & & \\ & 0 & \\ & & L \end{bmatrix} \begin{bmatrix} \dot{V}_1 \\ \dot{V}_2 \\ \dot{I} \end{bmatrix} = -\begin{bmatrix} G_{11} & G_{12} & -A_n^T \\ G_{21} & G_{22} & -A_n^T \\ A_n & A_{12} & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ I \end{bmatrix} + \begin{bmatrix} U \\ 0 \\ 0 \end{bmatrix} \qquad (4.1)$$

The first order derivative of intersection node voltage can be approximated as

$$\dot{V}_1(t+h) = -C^{-1}[G_{11}V_1(t) + G_{12}V_2(t) - A_n^T I(t) + U(t+h)] \qquad (4.2)$$

The inversion of capacitance matrix is easy because the capacitance matrix is actually a diagonal matrix.

Those nodes with relatively larger voltage derivative are selected as adaptive coarse nodes. We apply adaptive coarsening at the finest level. The coarse grid selection at other levels is determined solely by the coloring scheme, because active components with finer first level coarse grid will still have relatively finer grid on the next coarse levels. Figure 4.1 shows the comparison of non-adaptive and adaptive coarsening structure. Darker color means coarser level. The circuit activity focuses on the up-right corner of the circuits. Active region has finer grid than the other region in every grid level.
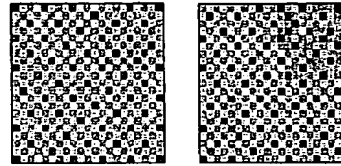

Figure 4.1 Non-Adaptive (left) and Adaptive (right) Coarsening Structure

4.2 Adaptive Smoothing Operation

To avoid the unnecessary computation for inactive regions, we reduce the global fine grid to several local fine grids corresponding to active regions at finest level and iterate the smoothing operator only inside these local fine grids. Although, in this paper, only finest grid is not global, we could have several levels of non-global grid. The active regions are detected by the first order derivative of nodal voltage (4.2).

With local fine grid, we actually use different time steps for active and inactive subnetworks in the sense that inactive subnetworks may only get chance to get error smoothed at finest level once every several time points.

5. EXPERIMENTAL RESULTS

The proposed approach is implemented in ANSI C. The experiments are executed on a SUN Blade100 (300MHz) workstation with 2GB memory.

We set the number of pre-smoothing and post-smoothing iteration to 3 at each level and the multigrid iteration termination control

residue norm to 1×10^{-4}. Gauss-Seidel method acts as smoothing operator at each level.

Table 1 shows the DC analysis runtime of SPICE3 and our approach, which is more than 100 times faster than SPICE3 for large circuits. Table 2 compares the transient analysis runtime of SPICE3, non-adaptive AMG, and adaptive AMG. The time-varying currents are modeled as triangular waveforms with 2mA peak current and 40ps rising and falling time [21]. Current sources are not evenly distributed and the timings are also different. We execute the transient analysis for 5ns. Experimental results show that adaptive AMG is about 10% faster than non-Adaptive AMG. Our approach runs more than 30 times faster than SPICE3 for transient analysis. The performance speedup is comparable with the PCG approach in [10]. The number of multigrid iterations dose not increase rapidly with the problem size. Actually the number of iterations is independent of the problem size, which cannot be proved but is observed in many cases.

Table 1 UC Analysis Runtime Comparison (sec)

Nodes	SPICE	AMG	Speedup
1706	1.68	0.17	9.8
2637	3.91	0.28	14
5105	15.01	0.57	26.3
10322	54.44	0.98	55.5
40842	708.22	3.93	180.2
91562	X	8.98	

Table 2. Transient Analysis Runtime Comparison (sec)

Nodes	SPICE	AMG	Adaptive AMG	Adaptive Speedup
1706	18.06	1.48	1.38	13.1
2637	41.23	4.22	3.92	10.5
5105	122.1	9.05	7.86	15.5
10322	456.42	19.34	16.97	26.9
40842	5048.5	96.5	87.55	57.7

Figure 5.1 compares the transient analysis voltage waveforms of one node from SPICE and our adaptive AMG approach. The waveforms are almost the same.
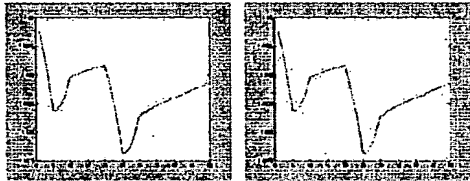

Fig 5.1 Comparison of voltage waveforms from SPICE (left) and Adaptive AMG (right)

6. CONCLUSION

In this paper, we present an adaptive AMG approach for power network analysis. The spatial variation of power supply noise is utilized to further speed up the performance. The new approach can handle power network of general topology with self and mutual inductances. Experimental result is very promising.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] S. Lin, N. Chang, "Challenges in Power-Ground Integrity", IEEE/ACM International Conference on Computer Aided Design. 2001.

[2] S. Bobba, T. Thorp, K. Aingaran, D. Liu, "IC power distribution challenges," IEEE/ACM International Conference on Computer Aided Design 2001

[3] J. R. Balck, "Electromigration Failure Modes in Aluminum Metalization for Semiconductor Devices", Proc. IEEE, pp. 1587-1594, Sep. 1969

[4] G. A. Katopis, "Delta-I Noise Specification for a High-performance Computing Machine," Proc. of the IEEE, vol. 73, pp.1405-1415, 1985

[5] S. Taylor, "The Challenge of Designing Global Signals in UDSM CMOS," IEEE Custom Integrated Circuits Conference, San Diego, CA 1999

[6] M. Zhao, R. V. Panda, S. S. Sapatnekar, D. Blaauw, "Hierarchical analysis of power distribution networks," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol.21, (no.2), IEEE, Feb. 2002. p.159-68.

[7] Y. Cao, Y. Lee, T. Chen, and C. Chen, "HiPRIME: Hierarchical and Passivity Reserved Interconnect Macromodeling Engine for RLKC Power Delivery", IEEE/ACM Design Automation Conference 2002

[8] H. Chen, J. Neely, "Interconnect and Circuit Modeling Techniques for Full-Chip Power Supply Noise Analysis", IEEE Transactions on Components, Packaging and Manufacturing Technology, Part B, Vol21, No.3, August 1998

[9] Y. Lee and C. Chen, "Power Grid Transient Simulation in Linear Time Based on Transmission-Line-Modeling Alternating-Direction-Implicit Method," IEEE/ACM International Conference on Computer Aided Design, 2001

[10] T. Chen and C. Chen, "Efficient Large-Scale Power Grid Analysis Based on Preconditioned Krylov-Subspace Iterative Methods", IEEE/ACM Design Automation Conference, 2001.

[11] S. R. Nassif et al, "Fast Power Grid Simulation", IEEE/ACM Design Automation Conference 2000.

[12] J. N. Kozhaya, S. R. Nassif, F. N. Najm, "Multigrid-like Technique for Power Grid Analysis", IEEE/ACM International Conference on Computer Aided Design. 2001.

[13] K. Stuben, "Algebraic Multigrid: An Introduction with Applications", GMD Report No.53, March 1999

[14] K. Stuben, "A review of algebraic multigrid" Journal of Computational and Applied Mathematics, vol.128, (no.1-2), Elsevier, 1 March 2001. p.281-309

[15] W. L. Briggs, "A Multigrid Tutorial", SIAM 2000

[16] "Design of high-performance microprocessor circuits," IEEE Press, 2001

[17] A. Devgan, .et al, "How to Efficiently Capture On-Chip Inductance Effects: Introducing a New Circuit Element K", IEEE/ACM International Conference on Computer Aided, Nov 2000, pp 150-155

[18] G. H. Golub, "Matrix Computations", Johns Hopkins University Press, 1996

[19] S. F. McCormick,, "Multilevel Adaptive Methods for Partial Differential Equations," vol.6, Frontiers in Applied Mathematics, SIAM, Philadelphia, 1989

[20] A. Brandt, "Multi-level adaptive solutions to boundary value problems," Math. Comput., 31, 1977, pp.333-390

[21] S. Zhao, K. Roy, C. K. Koh "Frequency domain analysis of switching noise on power supply network," IEEE/ACM International Conference on Computer Aided Design. 2000

ATTACHMENT II

Exemplary Functions and Applications:

The Spice$_{Diego}$ is a fast accurate transistor level simulation tool that can perform full-chip simulation including power network and clock network analysis with SPICE-level accuracy and guaranteed convergence. It can perform analysis of interconnect delay, crosstalk, voltage drop, ground bounce, simultaneous switching noise and EMI.

Exemplary Features and Advantages:

SPICE$_{Diego}$ adopts algebraic multigrid method as the solver engine, which avoids the high complexity of Gauss-Elimination method used in SPICE and slow convergence of basic iterative methods such as SOR and Gauss-Seidel.

SPICE$_{Diego}$ fully explores the circuit spatial and temporal latency as well as the multi-rate behavior to avoid unnecessary computation without loss of accuracy and convergence via the novel adaptive extension of multigrid method.

In addition, SPICE$_{Diego}$ has the ability to analyze coupling effect with electromagnetic retardations for high frequency chip-package-board scale analysis, while no commercial tools are available to consider the electromagnetic retardations.

Nassda's HSIM and Synopsys's PowerMill / TimeMill / RailMill are relied on circuit partitioning or macro modeling method which has potential convergence problem and cannot capture coupling effect such as mutual inductance accurately. Apache's Redhawk-SDL is actually cell-based. Simplex (now part of Cadence)'s VOLTAGESTORM ignores the inductance effect.

Section A: SPICE$_{Diego}$ Matrix Solver Engine

Multigrid Method

SPICE$_{Diego}$ adopts algebraic multigrid method as the solver engine. The multigrid method is an efficient technique first widely used for solving partial differential equations. The basic idea of multigrid method is to map the hard-to-damp low frequency error at fine level to easy-to-damp high frequency error at coarse level, solve the mapped problem at coarse level, and then map the error correction of coarse level back to fine level. A hierarchical grid structure is constructed for that purpose. At each level, a forward iterative smoothing operator such as Gauss-Seidel erases high frequency error. There are two kinds of multigrid methods: geometric multigrid and algebraic multigrid (AMG). Geometric multigrid method requires regular mesh structure. For problems with irregular structure, AMG is a good alternative of geometric multigrid method. The coarsening and interpolation operation of AMG are based on matrix itself. This overhead makes AMG less efficient than geometric multigrid if the problem analyzed has regular mesh structure. Since digital or mixed-signal circuits usually have irregular structures, SPICE$_{Diego}$ adopts the algebraic multigrid method. Another reason is that mutual inductance can be easily handled with AMG. Accuracy is guaranteed by comparing the norm of residue with user-defined error tolerance. The multigrid iteration stops only after the norm of residue is smaller than the error tolerance. Promising speedup over SPICE is observed in the experiments.

A description of single Multigrid cycle is shown below. The mapping operator from fine to coarse level is called restriction operator I_h^{2h}, coarse to fine level mapping operator is called interpolation operator I_{2h}^h. System equation $A^h u^h = f^h$ is solved hierarchically at different levels; the superscript refers to the level of the grid structure.

Single Multigrid-Cycle Scheme

$v^h \leftarrow V^h(v^h, f^h)$

- Relax on $A^h u^h = f^h$ v_1 times with initial guess v^h, calculate residue $r^h = f^h - A^h v^h$
- Compute $f^{2h} = I_h^{2h} r^h$
  - o Relax on $A^{2h} u^{2h} = f^{2h}$ v_1 times with initial guess $v^{2h} = 0$, $r^{2h} = f^{2h} - A^{2h} v^{2h}$
  - o Compute $f^{4h} = I_{2h}^{4h} r^{2h}$
    - Relax on $A^{4h} u^{4h} = f^{4h}$ v_1 times with initial guess $v^{4h} = 0$, $r^{4h} = f^{4h} - A^{4h} v^{4h}$
    - Compute $f^{8h} = I_{4h}^{8h} r^{4h}$
      - •
      - • to coarsest level
      - •
    - Correct $v^{4h} \leftarrow v^{4h} + I_{8h}^{4h} v^{8h}$
    - Relax on $A^{4h} u^{4h} = f^{4h}$ v_2 times with initial guess v^{4h}
  - o Correct $v^{2h} \leftarrow v^{2h} + I_{4h}^{2h} v^{4h}$
  - o Relax on $A^{2h} u^{2h} = f^{2h}$ v_2 times with initial guess v^{2h}.
- Correct $v^h \leftarrow v^h + I_{2h}^h v^{2h}$

Relax on $A^h u^h = f^h$ v_2 times with initial guess v^h

System Matrix Properties

For circuits with passive elements, when only RC model is included, system equation can be formulated as

$$C\dot{X}(t) + GX(t) = U(t),\tag{2.1}$$

where X is the vector of nodal voltages.

Appling trapezoidal approximation with time step h to equation (2.1), we have

$$(G + \frac{2}{h}C)X(t + h) = -(G - \frac{2}{h}C)X(t) + U(t) + U(t + h)\tag{2.2}$$

The left hand side matrix in (2.2) is symmetric and positive definite which makes the iterative methods converge quickly.

When inductance is included, the system matrix equation becomes:

$$\hat{C}\dot{\hat{X}}(t) + \hat{G}\,\hat{X}(t) = \hat{U}(t)\tag{2.3}$$

where $\hat{C} = \begin{bmatrix} C & 0 \\ 0 & L \end{bmatrix}$, $\hat{X} = \begin{bmatrix} V \\ I \end{bmatrix}$, $\hat{U} = \begin{bmatrix} U \\ 0 \end{bmatrix}$ and $\hat{G} = \begin{bmatrix} G & -A_l^T \\ A_l & 0 \end{bmatrix}$

Equation (2.3) can be rewritten as:

$$\begin{bmatrix} C & 0 \\ 0 & L \end{bmatrix}\begin{bmatrix} \dot{V} \\ \dot{I} \end{bmatrix} + \begin{bmatrix} G & -A_l^T \\ A_l & 0 \end{bmatrix}\begin{bmatrix} V \\ I \end{bmatrix} = \begin{bmatrix} U \\ 0 \end{bmatrix}\tag{2.4}$$

Appling trapezoidal approximation with time step h, the solution of equation (2.4) is derived by

$$\begin{bmatrix} \frac{2C}{h} + G & -A_l^T \\ A_l & \frac{2L}{h} \end{bmatrix}\begin{bmatrix} V(t+h) \\ I(t+h) \end{bmatrix} = \begin{bmatrix} \frac{2C}{h} - G & A_l^T \\ -A_l & \frac{2L}{h} \end{bmatrix}\begin{bmatrix} V(t) \\ I(t) \end{bmatrix} + \begin{bmatrix} U(t+h) + U(t) \\ 0 \end{bmatrix}\tag{2.5}$$

Although MNA method can handle elements without admittance description, the transient analysis system matrix in (2.5) is no longer symmetric and positive definite when the inductance is included because of the introduction of current variables.

Since Multigrid as well as PCG method requires the matrix to be symmetric positive definite, some extra processing is needed to reformulate the system. We split the variable vector into nodal voltage vector and branch current vector. Using block matrix operations, we decompose (2.5) into two iteration formulas (2.6) for nodal voltages and branch currents, respectively.

$$\left(\frac{2C}{h}+G-\frac{h}{2}A_l^T L^{-1}A_l\right)V(t+h)= \tag{2.6}$$
$$\Rightarrow\left\{\begin{array}{c}\left(\frac{2C}{h}-G+\frac{h}{2}A_l^T L^{-1}A_l\right)V(t)+[U(t)+U(t+h)]+2A_l^T I(T)\\[2mm] I(t+h)=I(t)-\frac{h}{2}L^{-1}A_l[V(t+h)+V(t)]\end{array}\right.$$

In equation (2.6), L^{-1} corresponds to the K matrix so that the matrix inversion overhead is reduced by sparsification methods.

If the inductance matrix is symmetric and positive definite (S.P.D.), we can prove that the system matrix is still S.P.D. This also holds for forward Euler, backward Euler integration approximation methods. Note that the topology of reformulated system matrix in (2.6) is no longer the same as original circuit topology. Geometric based coarse grid reduction algorithm cannot be applied to RLC network directly.

For circuits with nonlinear transistor devices, though the positive definite properties cannot be guaranteed, the system matrix usually exhibits this property especially for circuits with no strong feedback or high-gain elements. Thus, the fast AMG solver can be still be used.

Section B: Exploitation of Latency & Multi-rate Behavior

Most circuits display spatial and temporal latency as well as multi-rate behavior. Spatial latency means only part of circuits has activities at any time point. Temporal latency refers to the situation that a given portion of circuit is active in some periods but inactive in other periods. Even active portions have various change rate of current and voltage level. This phenomenon can be utilized to avoid unnecessary computations. Different regions can be simulated with various time step size according to their activities, while SPICE have to apply the minimum time step size to the entire circuits for the sake of convergence. Many partition-based commercial fast simulators use varies time step size for each subcircuit and solve each subcircuit separately. Though saving a lot of computation overhead, this kind of method is difficult to guarantee the convergence and cannot capture mutual inductance coupling effect correctly.

Instead of trying to directly employ varies time step sizes, SPICE$_{Diego}$ uses adaptive grid structures and adaptive smoothing operations to capture the multi-rate behavior and circuit latency. Active regions have relatively finer grid structure and more error smoothing operations than inactive regions. Convergence is guaranteed by the multigrid method, because multigrid converges as long as smoothing operations at each level can damp the high frequency error.

With the adaptive grid structure and adaptive smoothing, we actually use different "time step size" for active and inactive subcircuits in the sense that inactive subcircuits may only get chance to have error smoothed at finest level once every several time points as shown in the figure 3.1.
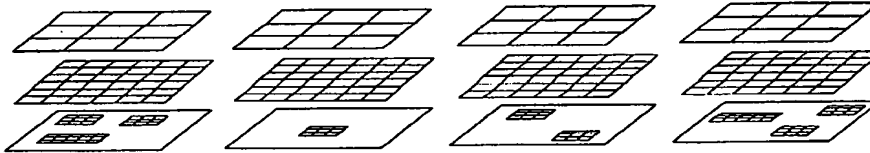

Figure 3.1 Adaptive Grid Structures

In SPICE$_{Diego}$, if node voltage change in the linear network (power or clock network) between time points is less than a threshold M, this node is considered as an idle node. If the change of Vgs, Vds and Vbs for a transistor is smaller than thresholds, the transistor is considered inactive. For active portions, different level of threshold could be defined to determine various degrees of activities that results in multi local fine levels of grid structures.

Section C: Device Model Evaluation

Nonlinear device model evaluation takes more than 60% of transient simulation runtime and thus a bottleneck of the circuit analysis. Many contemporary commercial simulation tools have an option to use simplified device models instead of detailed models such as BSIM3/BSIM4.

In the SPICE$_{Diego}$, a table MOSFET model is supported besides detailed BSIM3/BSIM4 models. The model table is generated from many detailed model evaluation runs before simulation. It is indexed by Vgs, Vds and Vbs and contains Ids, Gds (output conductance $Gds=\frac{\partial Ids}{\partial Vds}$) and Gm (Transconductance $Gm=\frac{\partial Ids}{\partial Vgs}$) as table items.

Experiments show that uniform voltage grid resolution of 60, 60 and 200 for Vgs, Vds and Vbs respectively and linear interpolation produce very accurate results compared with detailed models as shown in Figure 4.1.
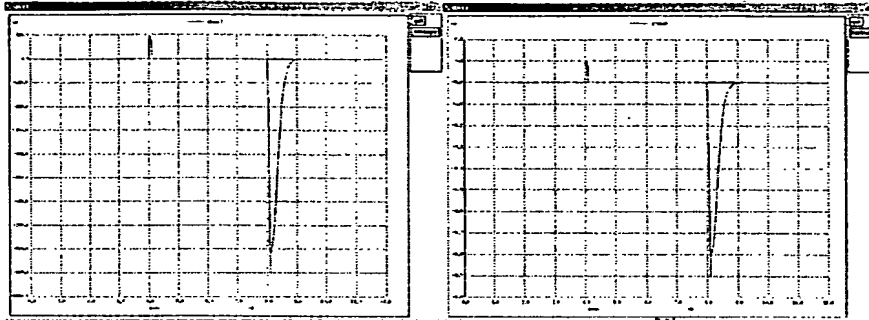

Figure 4.1 Current drawn from VDD of an inverter using BSIM3 Model (left) BSIM3 Table Model (right)
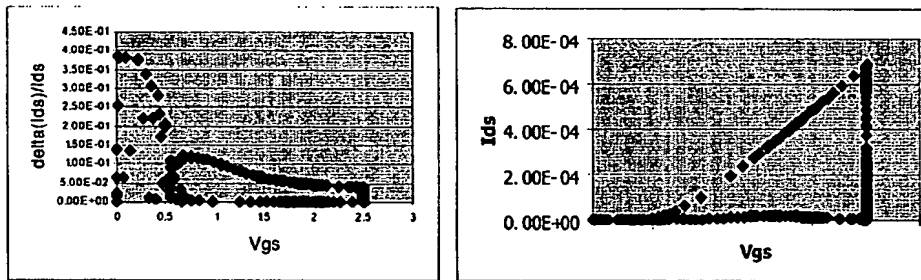

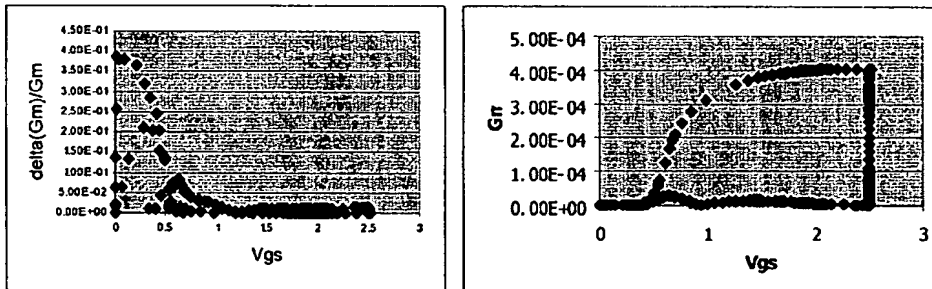**Figure 4.2 (a) Error Percentage of Ids vs Vgs (left) Ids vs Vgs (right)**

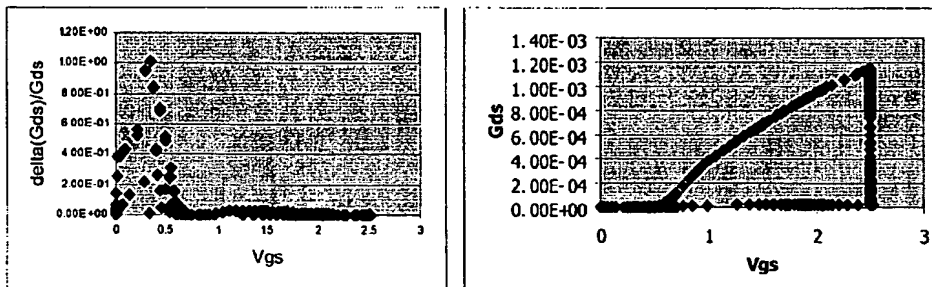**Figure 4.2 (b) Error Percentage of Gm vs Vgs (left) Gm vs Vgs (right)**

Figure 4.2 (c) Error Percentage of Gds vs Vgs (left) Gds vs Vgs (right)

Figure 4.2 shows the plots for the error percentage and absolute value of Ids, Gm and Gds vs Vgs of NMOS transistor in an inverter with a step function voltage at the gate terminal. The devices model is BSIM3v3 with TSMC .25 technology parameters. As observed in figure 4.2, table model is accurate when Ids/Vds/Gm is non-zero. Large errors at the left end of Vgs axis are mainly due to the close-to-zero absolute values of

Ids/Gm/Gds and hence can be ignored. Moreover, due to the non-uniform distributed first derivative value of Ids, Gm and Gds as shown in figure 4.3, non-uniform voltage grid resolution can improve the accuracy without requiring more memory.
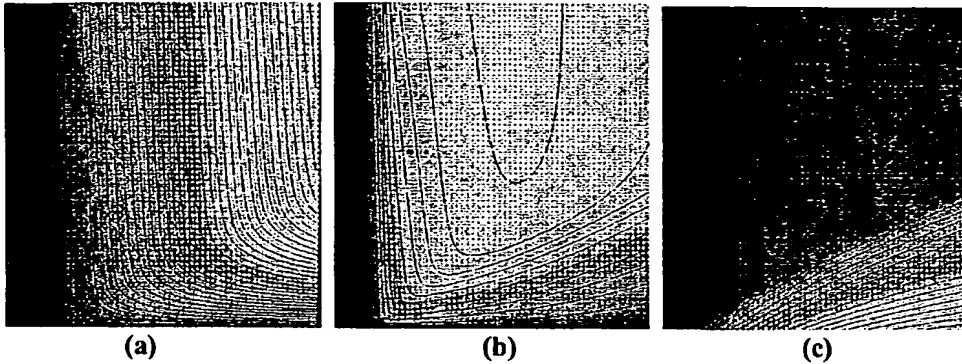

Figure 4.3 Contours of the (a) Ids (b) Gm (c) Gds with Vgs as X-axis and Vds as Y-axis

MOSFET capacitance has several options from detailed BSIM3/BSIM4 model to piecewise constant linear capacitances. 3-4 Newton-Raphson iterations are avoided for each device.

Most of the product features have been implemented in C on Unix platform. Some experiments have been conducted for power network analysis. Promising speedup has been observed in these experiments that are executed on a SUN Blade100 (300MHz) workstation with 2GB memory.

Table 1 shows the DC analysis runtime of SPICE3 and $SPICE_{Diego}$, which is more than 100 times faster than SPICE3 for large circuits. Table 2 compares the transient analysis runtime of SPICE3 and $SPICE_{Diego}$. We execute the transient analysis for 5ns. Experimental results show that $SPICE_{Diego}$ runs more than 30 times faster than SPICE3 for transient analysis. The speedup increases to 60 times when the size of the circuit grows to 40 thousand nodes. The number of multigrid iterations dose not increase rapidly with the problem size. Actually the number of iterations is independent of the problem size, which cannot be proved but is observed in many cases.

Table1 DC Analysis Runtime Comparison (sec)

Nodes	SPICE	$SPICE_{Diego}$	Speedup
1706	1.68	0.17	9.8
2637	3.91	0.28	14
5105	15.01	0.57	26.3
10322	54.44	0.98	55.5
40842	708.22	3.93	180.2
91562	X	8.98	

Table 2. Transient Analysis Runtime Comparison (sec)

Nodes	SPICE	$SPICE_{Diego}$	Speedup
1706	18.06	1.38	13.1
2637	41.23	3.92	10.5
5105	122.1	7.86	15.5
10322	456.42	16.97	26.9
40842	5048.5	87.55	57.7

Figure 5.1 compares the transient analysis voltage waveforms of one node from SPICE and $SPICE_{ucsd}$. The waveforms are almost the same.
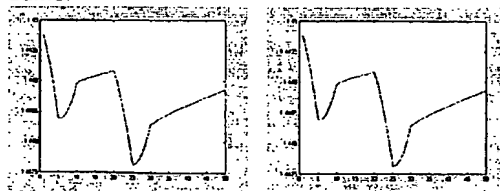
Figure 5.1 Comparison of voltage waveforms from SPICE (left) and SPICE_{Diego} (right)

<u>**Exemplary Commercial Applications**</u>

The Spice$_{Diego}$ is a fast accurate transistor level simulation tool that can perform full-chip simulation including power network and clock network analysis with SPICE-level accuracy and guaranteed convergence. Applications include but are not limited to analysis of interconnect delay, crosstalk, voltage drop, ground bounce, simultaneous switching noise and EMI, which is vital to chip designers.

10263139.doc